Exercise 1

1) Linear Independency and Gram-Schmidt process

Linearly independent?

To check if the columns of the matrix are linear independent we can compute the determinant of the matrix is non zero:

$$\det\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}] = [(1*2*3)+(2*3*0)+(0*0*3)] - [(0*2*3)+(0*2*3)+(0*3*1)]$$

$$= [6+0+0] - [0] = 6 \neq 0$$

So we have verified that the columns of the matrix are indeed linearly independent.

Gram-Schmidt process:

Used to obtain orthonormal vectors that are easier to work with but span the same space as the original column vectors

Take the first vector $\vec{v_1} = (1,0,0)^T$ and normalize it. But it is already normalized so we gonna refer to it as $\vec{u_1}$

Now we take the second vector $\vec{v_2} = (2,2,0)^T$ and compute its projection onto $\vec{u_1}$.

$$\text{proj}_{\vec{u_1}}(\vec{v_2}) = \frac{\vec{v_2}\cdot\vec{u_1}}{||\vec{u_1}||_2}\cdot\vec{u_1}$$

$$= \frac{1*2+2*0+0*0}{1} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

Next we take this projection and subtract it from $\vec{v_2}$

$$\vec{v_2}' = \vec{v_2} - \text{proj}_{\vec{u_1}}(\vec{v_2}) = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

Then we normalize it by dividing by its $||\ldots||_2$ Norm which is clearly 2 and we yield $\vec{u_2} = (0,1,0)^T$

For the third vector $\vec{v_3} = (3,3,3)^T$ we calculate its projection onto $\vec{u_1}$ and $\vec{u_2}$ and subtract both from $\vec{v_3}$:

$$\text{proj}_{\vec{u_1}}(\vec{v_3}) = \frac{\vec{v_3} \cdot \vec{u_1}}{||\vec{u_1}||_2} \cdot \vec{u_1}$$

$$= 3 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

and

$$\text{proj}_{\vec{u_2}}(\vec{v_3}) = \frac{\vec{v_3} \cdot \vec{u_2}}{||\vec{u_2}||_2} \cdot \vec{u_2}$$

$$= 3 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$$

We subtract it from $\vec{v_3}$:

$$\vec{v_3}' = \vec{v_3} - \text{proj}_{\vec{u_2}}(\vec{v_3}) - \text{proj}_{\vec{u_1}}(\vec{v_3}) = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

We normalize it by dividing by its norm (3) and yield the last orthonormal vector $\vec{u_3} = (0, 0, 1)^T$

So in total we get the set of vectors:

$$\{\vec{u_1}, \vec{u_2}, \vec{u_3}\} = \{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\}$$

Which are clearly orthonormal and process the same Span as the original vectors but are easier to work with

2) reflection matrix

Definition of the reflection matrix :

$$H = I - \frac{2}{w^*w} w w^*$$

a) H is self-adjoint (hermitian) and unitary/orthogonal and involutory

Hermitian
For that we compute H^*

$$H^* = \left(I - \underbrace{\frac{2}{\underbrace{w^*w}_{\in R}} \underbrace{ww^*}_{\in K^{n \times n}}} \right)^*$$

$$= I - \frac{2}{w^*w}(ww^*)^* = I - \frac{2}{w^*w}(w^*)^*w^*$$

$$= I - \frac{2}{w^*w}ww^* = H$$

unitary/orthogonal and **involutory**

For that we have to show $H^*H = I$. Since we already know that H is hermitian this simplifies to H^2.

$$H^2 = \left(I - \frac{2}{w^*w}ww^* \right) \cdot \left(I - \frac{2}{w^*w}ww^* \right)$$

$$= I^2 - 2 \cdot I \cdot \frac{2}{w^*w}ww^* + \left(\frac{2}{w^*w}ww^* \right)^2$$

$$= I^2 - 4\frac{ww^*}{w^*w} + 4\underbrace{\left(\frac{ww^*}{w^*w} \right)^2}_{\frac{w(w^*w)w^*}{(w^*w)^2} = \frac{ww^*}{w^*w}}$$

$$= I^2 = I$$

b) Determine eigenvalues, determinant and singular values

Luckily there is no need for computing the eigenvalues. Since the reflection matrix reflects across the subspace orthogonal to the vector w. So any vector in the direction of w will be "flipped" with a eigenvalue of -1. Any vector in a orthogonal subspace to w will remain unchanged with an eigenvalue $+1$.

$$\lambda_1 = -1 \quad \text{(with multiplicity 1)}$$

\lambda_2 = 1 \quad \text{(with multiplicity n-1)},

The determinant can be computed by calculating the product of the eigenvalues.

$$\det(H) = \prod_i^n \lambda_i = -1 \cdot 1^{(n-1)} = -1$$

The singular values of a matrix H are the square roots of the eigenvalues of the matrix H^*H. Since we already showed that H is **unitary** (or **orthogonal**), we have $H^*H = I$. This means all eigenvalues of H^*H are 1, and thus the singular values of H are all 1.

Therefore, the **singular values of H** are:

$$\sigma_1 = \sigma_2 = \cdots = \sigma_n = 1$$

3) Norm (in)equalities

$**||x||_\infty \le ||x||_2$

The infinity norm picks the largest element of $x \in K^n$ while the Euclidean norm

$|x||2 = \left(\sum i = 1^n |x_i|^2 \right)^{1/2}$

Therefore the $||x||_2$ is always as least as large as the $||x||_\infty$ norm.

Equality holds if all the elements have the same value e.g: $(1, 0, 0, \dots, 0)^T$

$||x||_2 \le \sqrt{n}||x||_\infty$

This holds because the $||x||_\infty$ norm gives us an upper bound on the elements that get squared and summed. So for sure $n||x||_\infty \ge \sum_{i=1}^{n} x_i$. If we then look at the sum of squares:

$$n||x||_\infty \ge \sum_{i=1}^{n} x_i^2$$

$$\sqrt{n}||x||_\infty \ge \sqrt{\sum_{i=1}^{n} x_i^2}$$

Equality will hold if all elements are the same e.g: $(1, 1, \dots, 1)^T$

4) Norm limits

The Definition of the p-Norm is:

$$||x||_p = \left(\sum_{i}^{n} |x_i|^p \right)^{1/p}$$

Intuition

The larger p gets the more the sum will get dominated by the p-th power of the largest term which is by definition the $||x||_\infty$. Lets assume $||x||_\infty = x_k$

So for very large p:

$$||x||_p = \left(\sum_{i}^{n} |x_i|^p \right)^{1/p} \approx x_k$$

And as p gets bigger and bigger the domination of the largest term will become stronger and stronger. Therefore taking the limit does indeed converge to the largest element or in other words the ∞-Norm.

5) Frobenius Norm $||A||_F$

We have to show that:

$$||A||_F = (\text{Tr}(A^*A))^{1/2} = (\text{Tr}(AA^*))^{1/2}$$

With $||A||_F$ being defined as follows:

$$||A||_F = \sqrt{\sum_i^d \sum_k^d |A_{ik}|^2}$$

In general the trace has the following property : $\text{Tr}(AB) = \text{Tr}(BA)$ and is defined as the sum over the diagonal entries $\text{Tr}A = \sum_i A_{ii}$

Now lets look at A^*A more closely:
The conjugate transpose works as follows: $(A_{ij})^* = \bar{A}_{ji}$

$$(A^*A)_{ij} = \sum_{k=1}^d (A_{ik})^* A_{kj} = \sum_{k=1}^d A_{ki} A_{kj}$$

The diagonal entries are $(A^*A)_{ii} = \sum_{k=1}^d (A_{ik})^* A_{ki} = \sum_{k=1}^d |A_{ki}|^2$

Putting this all together the trace becomes:

$$\text{Tr}(A^*A) = \sum_{i=1}^d (A^*A)_{ii}$$

$$= \sum_{i=1}^d \sum_{k=1}^d |A_{ik}|^2$$

So we have shown that $||A||_F^2 = \sum_{i=1}^d \sum_{k=1}^d |A_{ik}|^2$. A trivial application of the sqare root will yield the desired result.
In principle the Tr operation is invariant under cyclic permutations. But the proof for the second equality would follow the same outline as seen above

6) Condition numbers κ_1, κ_2 and κ_∞

The **condition number** of a matrix measures how sensitive the solution of a linear system is to small changes or errors in the input. A high condition number indicates that even small errors in the input (such as round-off errors) could lead to large errors in the solution. The condition number depends on the matrix norm being used.

The definition of the p-kondition number is:

$$\kappa_p = ||A||_p|| \cdot ||A^{-1}||_p$$

So lets first compute the inverse of the matrix. A quick check shows that the determinant of the matrix is not zero.

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$||A||_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |A_{ij}|$$

which is simply the maximum absolute column sum of the matrix.

$$\kappa_1 = ||A||_1|| \cdot ||A^{-1}||_1 = 3 \cdot 1 = 3$$

$p = 2$

$$||A||_2 = \sqrt{\lambda_{max}(A^*A)} = \sigma_{max}(A)$$

The singular values of a matrix H are the square roots of the eigenvalues of the matrix $H^* H$.

So lets calculate A^*. Which is not difficult since all values are real. So all we have to do is to transpose the matrix.

$$A^* = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = A$$

Which means that the matrix is hermitian. So $A^*A = A^2$

$$A^2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}$$

Let's calculate the eigenvalues by solving $\det(A^2 - I\lambda)$

$$\det(A^2 - I\lambda) = \det \begin{bmatrix} 5 - \lambda & -4 \\ -4 & 5 - \lambda \end{bmatrix} = (5 - \lambda)^2 - 16 = 0$$
$$= 5 - \lambda = \pm 4$$

So we find two solutions

$$\lambda = \begin{cases} \lambda_1 = 1 \\ \lambda_2 = 9 \end{cases}$$

So we find that the $||A||_2 = \sqrt{9} = 3$

$p = \infty$

$$||A||_\infty = \max_{1 \le i \le m} \sum_{i=1}^{n} |A_{ij}|$$

which is simply the maximum absolute row sum of the matrix.

So we get:

$$\kappa_\infty = ||A||_\infty \cdot ||A^{-1}||_\infty$$
$$= 3 \cdot 1 = 3$$

7) Conditioning and perturbation

We solve the unperturbed system $Ax = b$

unperturbed system

First we bring the matrix in upper triangle form:

$$\begin{bmatrix} 1 & 4 \\ 1 & 4.001 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 \\ 0 & 0.001 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

Which yields the following two equations:

$$x_1 + 4x_2 = 4$$
$$0.001x_2 = 0$$

So we have the solution:

$$x = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

perturbed system

$$\begin{bmatrix} 1 & 4 \\ 1 & 4.001 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0.001 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 \\ 0 & 0.001 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 0.001 \end{bmatrix}$$

So we have the following system of equations:

$$x_1 + 4x_2 = 4$$
$$0.001x_2 = 0.001$$

Which yields the solution

$$x' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

relative error

There is some theorem that all norms are equivalent so we are gonna use the $||\ldots||_2$ norm

$$\mathcal{E}_{rel} = \frac{||x' - x||_2}{||x||_2}$$
$$= \frac{\sqrt{(-4)^2 + 1^2}}{\sqrt{16}} = \frac{\sqrt{17}}{4} \approx 1.03$$

This means that the change in the solution is of the same magnitude as the original solution. Therefore the problem/linear-system is ill-conditioned.